International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

On the Optimal Transmission of Non-Gaussian Signals through a Noisy Channel with Feedback

Takeuchi, Y.

IIASA Working Paper

WP-91-002

January 1991

Working Paper

On the Optimal Transmission of Non-Gaussian Signals through a Noisy Channel with Feedback

*Yoshiki Takeuchi**

WP-91-2
January 1991

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria

Telephone: (0 22 36) 715 21*0 □ Telex: 079 137 iiasa a □ Telefax: (0 22 36) 71313

On the Optimal Transmission of Non-Gaussian Signals through a Noisy Channel with Feedback

*Yoshiki Takeuchi**

WP-91-2
January 1991

* On leave from the Department of Industrial Engineering,
Kanazawa Institute of Technology, 7-1 Ohgigaoka,
Nonoichi-machi, Ishikawa 921, Japan

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria

Telephone: (0 22 36) 715 21 *0 □ Telex: 079 137 iiasa a □ Telefax: (0 22 36) 71313

Preface

This paper is concerned with the optimal transmission of a non-Gaussian signal (a non-Gaussian message) through a channel with Gaussian white noise by a cording which is linear in the signal. Under the assumptions of square integrability on the signal and the independence between the signal and the noise, it will be shown that the optimal cording which maximizes the mutual information between the signal (the non-Gaussian message) and the observation process (the channel output) is to generate the estimation error process multiplied by a deterministic coefficient so that the mean power of the encorded signal takes the maximum admissible value. The result shows that the optimal transmission is such that the channel output becomes the innovations process.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

1. Introduction

In ralation with the linear and nonlinear filtering problems, there are a number of researches on the problem of the optimal transmission of stochastic processes through a Gaussian channel with feedback[1] - [8]. Most of them are concerned with Gaussian signals (i.e., Gaussian messages). Ihara[7] considered the optimal cording of a Gaussian process for transmission through a channel with Gaussian white noise. He showed that the optimal cording which maximizes the mutual information between the signal (the Gaussian message) and the observation process (the channel output) under a constraint on the mean power of the encorded signal is given by a functional which is linear in the signal. His result also shows that the optimal cording is composed of the two steps:

(1) minimization of the mean power of the encorded signal over the cordings with the same mutual information which is achieved by generating the estimation error process, i.e., the difference between the original signal and its optimal estimate.

(2) maximization of the mutual information by a multiplication by a deterministic coefficient which increases exponentially in time and is determined in such a way that the mean power of the encorded signal takes the maximum admissible value.

For the case of the Gaussian message generated by a linear stochastic differential equation, an explicit formula for the optimal transmission was obtained by Liptser and Shiryayev[8] in relation with conditionally Gaussian nonlinear filtering problem.

In this paper, we are concerned with the optimal transmission of a non-Guassian signal (a non-Gaussian message) through a channel with Gaussian white noise by a cording which is linear in the signal. Under the assumptions of the square integrability on the signal and the independence between the signal and the noise, it will be shown that the optimal transmission is described by a formula similar to the one in the Gassian case. It will also be seen that the optimal transmission is such that the channel output is the innovations process.

In this paper, mathematical symbols are used in the following way. The prime denotes the transpose of a vector or a matrix. The Euclidean norm is $|\cdot|$. If A is a nonsingular square matrix, A^{-1} denotes the inverse matrix of A. The triplet $(\Omega, \mathcal{F}, P)$ is a complete probability space where Ω is a sample space with elementary events ω, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and P is a probability measure. $E\{\cdot\}$ denotes the expectation and $E\{\cdot|\mathcal{G}\}$, $\mathcal{G}\subset\mathcal{F}$ the conditional expectation, given $\mathcal{G}$, with respect to P. $\sigma\{\cdot\}$ is the minimal sub-σ-algebra of $\mathcal{F}$ with respect to which the family of $\mathcal{F}$-measurable sets or random variables $\{\cdot\}$ is measurable. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are sub-σ-algebras of $\mathcal{F}$, then $\mathcal{F}_1 \vee \mathcal{F}_2$ denotes the minimal σ-algebra which contains both $\mathcal{F}_1$ and $\mathcal{F}_2$. Also, for $\mathcal{G}\subset\mathcal{F}$ and $A\in\mathcal{F}$, $\mathcal{G}\cap A$ denotes the family $\{B\cap A,\ B\in\mathcal{G}\}$. Let $F \equiv \{\mathcal{F}_t;\ 0 \le t \le T\}$ be a non-decreasing family of σ-algebras. A stochastic process $\mathbf{x} \equiv \{x_t;\ 0 \le t \le T\}$ is said to be *adapted to* F or F-*adapted* if $x_t(\omega)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$. It is assumed that all random variables and stochastic processes are $\mathcal{F}$-measurable. Unless otherwise stated, stochastic properties are that with respect to P.

2. The Optimal Transmission Problem

Let $\mathbf{f} \equiv \{f_t(\omega); 0 \le t \le T\}$ be an m-dimensional stochastic process which denotes the signal of our interest and satisfies

$$E\{\int_0^T |f_t(\omega)|^2 dt\} < \infty. \tag{1}$$

If $\mathbf{f}$ itself is sent through a channel with additive Gaussian noise, then the observation process to be received is represented by

$$\eta_t = f_t(\omega) + \dot{w}_t, \qquad t \in [0,T], \tag{2}$$

or equivalently,

$$y_t^o = \int_0^t f_s(\omega)\,ds + w_t, \qquad t \in [0,T], \tag{3}$$

where $\mathbf{w} \equiv \{w_t; 0 \le t \le T\}$ is an m-dimensional standard Brownian motion process independent of $\mathbf{f}$, and hence, $\dot{\mathbf{w}} \equiv \{\dot{w}_t; 0 \le t \le T\}$ is a white Gaussian noise independent of $\mathbf{f}$. For the observation (2) and/or (3), the mutual information between $\mathbf{f}$ and $\mathbf{y}^o$ is given by[6],[8]

$$I_t(\mathbf{f},\mathbf{y}^o) = \frac{1}{2}\int_0^t E\{|f_s(\omega) - \hat{f}_s^o|^2\}ds, \qquad t \in [0,T], \tag{4}$$

where

$$\hat{f}_t^o \triangleq E\{f_t(\omega)|\mathcal{Y}_t^o\}, \qquad t \in [0,T], \tag{5}$$

and

$$\mathcal{Y}_t^o \triangleq \sigma\{y_t^o; 0 \le s \le t\}, \qquad t \in [0,T]. \tag{6}$$

Also, the mean power of the signal is given by

$$P_t^o \triangleq \frac{1}{t} E\{\int_0^t |f_s(\omega)|^2 ds\}, \qquad t \in [0,T]. \tag{7}$$

Now, let us consider the problem of improving the efficiency of the transmission by using noiseless feedback[7],[8]. Let $\mathbf{y} \equiv \{y_t; 0 \le t \le T\}$ be the observation by the transmission with noiseless feedback described by

$$y_t = \int_0^t \beta_s(\mathbf{f},\mathbf{y})\,ds + w_t, \qquad t \in [0,T], \tag{8}$$

where $\beta_t(\mathbf{f},\mathbf{y}), 0 \le t \le T$ is a nonanticipative functional of $\mathbf{f}$ and $\mathbf{y}$ for which there exists a unique strong solution of (8). As it is well-known[8], for any solution $\mathbf{y}$ of (8), the mutual information between $\mathbf{f}$ and $\mathbf{y}$ is given by

$$I_t(\mathbf{f},\mathbf{y}) = \frac{1}{2}\int_0^t E\{|\beta_s(\mathbf{f},\mathbf{y}) - \hat{\beta}_s|^2\}ds, \tag{9}$$

where

$$\hat{\beta}_t \triangleq E\{\beta_t(\mathbf{f},\mathbf{y})|\mathcal{Y}_t\}. \tag{10}$$

and

$$\mathcal{Y}_t \triangleq \sigma\{y_s;\ 0 \leq t \leq T\}. \tag{11}$$

Then, we want to select a functional $\boldsymbol{\beta}$ in such a way that the mutual information given by (9) is maximized under the power constraint:

$$P_t \triangleq \frac{1}{t}\int_0^t E\{|\beta_s(\mathbf{f},\mathbf{y})|^2\}ds \leq P_t^o. \tag{12}$$

In this paper, we are concerned with the special case of this problem in which $\boldsymbol{\beta}$ is selected over a sub-class of functionals with the form:

$$\beta_t(\mathbf{f},\mathbf{y}) = H(t)f_t(\omega) - \phi(t,\mathbf{y}), \qquad t \in [0,T], \tag{13}$$

where $\mathbf{H} \equiv \{H(t);\ 0 \leq t \leq T\}$ is a deterministic time function and $\phi(t,\mathbf{y})$ is a nonanticipative functional of $\mathbf{y}$. As it is well-known, in the case where $\mathbf{f}$ is a Gaussian process with $E\{\int_0^T |f_t(\omega)|^2 dt\} < \infty$ which is independent of $\mathbf{w}$, the optimal functional $\boldsymbol{\beta}^*$ over the linear class given by (13) is also optimal over the class of all nonanticipative functionals for which (8) has a unique strong solution. The result[7] shows that the solution is given by the following mini-max scheme.

[*MINI-MAX SCHEME*] Let $\bar{\mathbf{y}}$ denote the process given by (8) with (13) for the case $\phi \equiv 0$, i.e.,

$$\bar{y}_t = \int_0^t H(s)f_s(\omega)ds + w_t, \qquad t \in [0,T]. \tag{14}$$

Also, let

$$\bar{f}_t \triangleq E\{f_t(\omega)|\bar{\mathcal{Y}}_t\}, \tag{15}$$

where

$$\bar{\mathcal{Y}}_t \triangleq \sigma\{\bar{y}_s;\ 0 \leq s \leq t\}. \tag{16}$$

(Step 1) For any $\mathbf{H} \equiv \{H(t); 0 \le t \le T\}$ satisfying

$$I_t(\mathbf{f},\bar{\mathbf{y}}) = \frac{1}{2}\int_0^t E\{|H(s)[f_s(\omega) - \bar{f}_s]|^2\}ds \le \frac{t}{2}P_t^o, \tag{17}$$

P_t is *minimized* by a functional ϕ for which there exists a unique strong solution, denoted by $\mathbf{y}^*$, of (8) with (13), and ϕ and $\mathbf{y}^*$ satisfy

$$\phi(t,\mathbf{y}^*) = H(t)\hat{f}_t^*, \quad \text{P-a.s.}, \quad t \in [0,T], \tag{18}$$

where $\hat{f}_t^* \triangleq E\{f_t(\omega)|\mathcal{Y}_t^*\}$ and $\mathcal{Y}_t^* \triangleq \sigma\{y_s^*; 0 \le s \le t\}$.

(Step 2) For any ϕ obtained by Step 1, we have

$$I_t(\mathbf{f},\mathbf{y}^*) = \frac{1}{2}\int_0^t E\{|H(s)[f_s(\omega) - \hat{f}_s^*]|^2\}ds = \frac{t}{2}P_t \ (\le \frac{t}{2}P_t^o) \tag{19}$$

Then, $I_t(\mathbf{f},\mathbf{y}^*)$ is *maximized* by choosing $\mathbf{H}$ so as to satisfy

$$E\{|H(t)[f_t(\omega) - \hat{f}_t^*]|^2\} = p(t) \triangleq E\{|f_t(\omega)|^2\}. \tag{20}$$

□

For the Gaussian case, the above mini-max scheme is valid because

(i) the functional with property (18) is linear and hence, there exists a unique strong solution of (8) with (13).

(ii) we have $I_t(\mathbf{f},\mathbf{y}^*) = I_t(\mathbf{f},\bar{\mathbf{y}})$, $0 \le t \le T$ because the conditional distribution of $\mathbf{f}$ given $\mathcal{Y}_t^*$ is conditionally Gaussian and is equal to the one given $\bar{\mathcal{Y}}_t$ [8].

Thus, according to the above mini-max scheme, we can call ϕ and $\mathbf{H}$ respectively "power coefficient" and "information coefficient" because ϕ is chosen to minimize the power and does not change the mutual information whereas $\mathbf{H}$ is chosen to maximize the mutual information and which implies that the power does not change but takes the preassigned value.

Example 1 (Case of Gaussian signals given by linear stochastic differential equations) [8]. Let us consider the optimal transmission problem for the case $\mathbf{f} \equiv \mathbf{x}$ where $\mathbf{x} \equiv \{x_t; 0 \le t \le T\}$ is an m-dimensional Gaussian process determined by

$$x_t = x_0 + \int_0^t A(s)x_s ds + \int_0^t G(s)d\bar{w}_s, \qquad t \in [0,T], \tag{21}$$

and where x_0 is an m-dimensional random variable with $E\{|x_0|^2\} < \infty$ and $\bar{\mathbf{w}} \equiv \{\bar{w}_t;\ 0 \le t \le T\}$ is a d-dimensional Brownian motion process. We will assume that x_0, $\bar{\mathbf{w}}$ and $\mathbf{w}$ are mutually independent. Let $\bar{\sigma}(t)$ denote the maximum signal intensity per component given by

$$\bar{\sigma}(t) \triangleq \sqrt{p(t)/m}\ . \tag{22}$$

Then, the optimal coefficients ϕ and $\mathbf{H}$ of (13) are given by

$$\phi(t,\mathbf{y}^*) = H(t)\hat{x}_t^* \tag{23}$$

and

$$H(t) = \bar{\sigma}(t)Q^{-\frac{1}{2}}(t), \tag{24}$$

where $\hat{x}_t^*$ and $Q(t)$ are given by

$$\begin{cases} d\hat{x}_t^* = A(t)\hat{x}^*dt + \bar{\sigma}(t)Q^{\frac{1}{2}}(t)\,dy_t^*, \\ \hat{x}_0^* = E\{x_0\}, \qquad t \in [0,T], \end{cases} \tag{25}$$

$$\begin{cases} Q(t) = \Phi(t,0)Q_0\Phi'(t,0) + \int_0^t \Phi(t,s)G(s)G'(s)\Phi'(t,s)\,ds, \\ Q_0 = E\{[x_0 - \hat{x}_0^*][x_0 - \hat{x}_0^*]'\}, \qquad t \in [0,T], \end{cases} \tag{26}$$

and $\Phi(t,s)$ is the solution of linear matrix differential equation:

$$\begin{cases} \dfrac{d\Phi(t,s)}{dt} = [A(t) - \dfrac{\bar{\sigma}^2(t)}{2}\cdot I]\Phi(t,s) \\ \Phi(s,s) = I, \qquad 0 \le s \le t \le T. \end{cases} \tag{27}$$

Consequently, the optimal transmission is described by

$$y_t^* = \int_0^t \bar{\sigma}(s)Q^{-\frac{1}{2}}(s)\{x_s - \hat{x}_s^*\}ds + w_t, \tag{28}$$

with (25) - (27) and (22). $\square$

It should be noted that in Example 1, the functional ϕ given by (23) is admissible because, as we can see from (25), $\hat{x}^*$ is a linear functional of $\mathbf{y}^*$ and therefore, (28) has a unique strong solution.

In this paper, we will show that the above mini-max scheme remains valid for a class of non-Gaussian signals.

3. Main Results

Let us assume that the following conditions are satisfied.

(C - 1) $E\{\int_0^T |f_t(\omega)|^2 dt\} < \infty$.

(C - 2) $\mathbf{f}$ and $\mathbf{w}$ are mutually independent.

(C - 3) For $\bar{\mathbf{y}}$ given by (14), there exists a function $\mathbf{H}$ which satisfies

$$I_t(\mathbf{f},\bar{\mathbf{y}}) = \frac{t}{2} P_t^o, \qquad t \in [0,T], \tag{29}$$

and

$$\sup_{0 \le t \le T} |H(t)| < \infty. \tag{30}$$

Then, the mini-max scheme given in the previous section remains valid, i.e.,

Theorem 1. For a non-Gaussian signal denoted by $\mathbf{f}$ and the observation described by (8) and (13), assume (C - 1) - (C - 3). Then, for any $\mathbf{H}$ satisfying (17) and (30), (18) determines a unique functional ϕ which is admissible and optimal in the sense:

(i) There exists a unique strong solution, $\mathbf{y}^*$, of (8) with (13).

(ii) The mutual information is unchanged by using ϕ, i.e.,

$$I_t(\mathbf{f},\mathbf{y}^*) = I_t(\mathbf{f},\bar{\mathbf{y}}), \qquad t \in [0,T], \tag{31}$$

where $\bar{\mathbf{y}}$ is given by (14), namely, the process given by (8) with (13) and $\phi \equiv 0$.

(iii) P_t, $0 \le t \le T$ given by (12) is minimized. $\square$

Let $\mathbf{v} \equiv \{v_t;\, 0 \le t \le T\}$ denotes the innovations process for $\bar{\mathbf{y}}$ defined by

$$v_t \triangleq \bar{y}_t - \int_0^t H(s)\bar{f}_s ds, \qquad t \in [0,T]. \tag{32}$$

Then, a more practical description of Theorem 1 is

Theorem 2. Under the assumptions of Theorem 1, the optimal functional ϕ, for any $\mathbf{H}$ satisfying (17) and (30), is given by

$$\phi(t,\mathbf{v}) = H(t)\bar{f}_t, \quad \text{P-a.s.}, \quad t \in [0,T], \tag{33}$$

and consequently, we have

$$\mathbf{y}^* = \mathbf{v}, \quad \text{P-a.s.}, \quad t \in [0,T], \tag{34}$$

i.e., the optimal transmission is realized by sending the innovations process. □

Proof. See Remark 3 below. □

Remark 1. As we can see from (17), the condition given by (30) is implied by

$$\inf_{0\le t\le T} E\{|f_t(\omega) - \bar{f}_t|^2\} > 0. \quad \square$$

Remark 2. The condition given by (31) is equivalent to

$$H(t)\hat{f}_t^* = H(t)\bar{f}_t, \quad \text{P-a.s.}, \quad t \in [0,T]. \tag{35}$$

This is easily seen as follows. Since

$$y_t^* = \int_0^t H(s) f_s(\omega)\,ds - \int_0^t \phi(s,\mathbf{y}^*)\,ds + w_t = \bar{y}_t - \int_0^t \phi(s,\mathbf{y}^*)\,ds,$$

we have

$$\bar{y}_t = y_t^* + \int_0^t \phi(s,\mathbf{y}^*)\,ds, \tag{36}$$

and which implies $\bar{\mathcal{Y}}_t \subset \mathcal{Y}_t^*$, $0 \le t \le T$. Hence, from (17), we have

$$\begin{aligned} I_t(\mathbf{f},\bar{\mathbf{y}}) &= \frac{1}{2}\int_0^t E\{|H(s)[f_s(\omega) - \hat{f}_s^*]|^2\}ds + \frac{1}{2}\int_0^t E\{|H(s)[\hat{f}_s^* - \bar{f}_s]|^2\}ds \\ &= I_t(\mathbf{f},\mathbf{y}^*) + \frac{1}{2}\int_0^t E\{|H(s)\hat{f}_s^* - H(s)\bar{f}_s|^2\}ds. \end{aligned} \tag{37}$$

Thus, we see that the conditions given by (31) and (35) are equivalent. □

Remark 3. Without proving Theorem 1, we can see that the functional ϕ with the properties (18) and (ii) of Theorem 1 is unique and is given by (33). This is easily seen by noting that, because of (18) and (35), we have

$$y_t^* = \bar{y}_t - \int_0^t \phi(s,\mathbf{y}^*)\,ds = \bar{y}_t - \int_0^t H(s)\hat{f}_s^* ds = \bar{y}_t - \int_0^t H(s)\bar{f}_s ds = v_t,$$

and hence,

$$\phi(t,\mathbf{v}) = \phi(t,\mathbf{y}^*) = H(t)\hat{f}_t^* = H(t)\bar{f}_t, \quad \text{P-a.s.} \tag{38}$$

The existence of the functional which satisfies (33) is guranteed by the innovations informational equivalence[9], i.e., $\bar{\mathcal{Y}}_t = \mathcal{V}_t$, $0 \leq t \leq T$, where $\mathcal{V}_t \triangleq \sigma\{v_s;\, 0 \leq s \leq t\}$. However, in order that ϕ, given by (33), becomes an admissible functional, we have to show (i). For, there might exist another solution of (8) which does not satisfy (18) and/or (ii). □

4. Proof of Theorem 1

For any $\mathbf{H}$ satisfying (17) and (30), let, for simplicity,

$$h_t(\omega) \triangleq H(t) f_t(\omega), \qquad t \in [0,T]. \tag{39}$$

Then, we can write

$$\bar{y}_t = \int_0^t h_s(\omega) ds + w_t, \qquad t \in [0,T], \tag{40}$$

$$y_t = \int_0^t \{h_s(\omega) - \phi(s,\mathbf{y})\} ds + w_t, \qquad t \in [0,T], \tag{41}$$

and

$$v_t = \bar{y}_t - \int_0^t \bar{h}_s ds = \int_0^t \{h_s(\omega) - \bar{h}_s\} ds + w_t, \qquad t \in [0,T], \tag{42}$$

where

$$\bar{h}_t \triangleq E\{h_t(\omega) | \bar{\mathcal{Y}}_t\}. \tag{43}$$

Now, let us start with finding a class of functionals ϕ with property (18), which is equivalent to

$$\phi(t,\mathbf{y}^*) = \hat{h}_t^*, \qquad \text{P-a.s.}, \qquad t \in [0,T], \tag{44}$$

where

$$\hat{h}_t^* \triangleq E\{h_t(\omega) | \mathcal{Y}_t^*\}. \tag{45}$$

On the probability space $(\Omega, \mathcal{F}, P)$, let us consider the equation:

$$\rho_t(\omega) = \int_{\tilde{\Omega}} h_t(\tilde{\omega}) \exp\{\int_0^t [h_s(\tilde{\omega}) - \rho_s(\omega)]' dv_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \rho_s(\omega)|^2 ds\} dP(\tilde{\omega}),$$
$$t \in [0,T], \tag{46}$$

where $\tilde{\Omega} \triangleq \Omega$ and $\tilde{\omega} \in \tilde{\Omega}$. We will define the solution of (46) as in the same way as usual stochastic differential equations[10],[11].

Definition 1 (Weak solutions). Let $\tilde{\mathbf{F}} \equiv \{\tilde{\mathcal{F}}_t;\ 0 \le t \le T\}$ denote any non-decreasing family of sub-σ-algebras of $\mathcal{F}$ and $\tilde{\mathbf{w}} \equiv \{\tilde{w}_t;\ 0 \le t \le T\}$ be any m-dimensional standard Brownian motion process. The triplet $(\boldsymbol{\rho},\tilde{\mathbf{w}},\tilde{\mathbf{F}}) \equiv \{(\rho_t(\omega),\tilde{w}_t,\tilde{\mathcal{F}}_t);\ 0 \le t \le T\}$, where $\boldsymbol{\rho} \equiv \{\rho_t(\omega);\ 0 \le t \le T\}$ is an m-dimensional stochastic process, is called *a (weak) solution* of (46) if the following conditions are satisfied.

(i) $\tilde{\mathbf{w}}$ is an $\tilde{\mathbf{F}}$-Brownian motion process.

(ii) $\boldsymbol{\rho}$ is adapted to $\tilde{\mathbf{F}}$.

(iii) $P\{\omega;\ \int_0^T |\rho_t(\omega)|^2 dt < \infty\} = 1$.

(iv) For all $t \in [0,T]$,

$$\rho_t(\omega) = \int_{\tilde{\Omega}} h_t(\tilde{\omega}) \exp\{\int_0^t [h_s(\tilde{\omega}) - \rho_s(\omega)]' d\tilde{w}_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \rho_s(\omega)|^2 ds\} dP(\tilde{\omega}),$$

P-a.s. (47) □

Definition 2 (Strong solutions). An m-dimensional stochastic process $\boldsymbol{\rho} \equiv \{\rho_t(\omega);\ 0 \le t \le T\}$ is called *a strong solution* of (46) if $(\boldsymbol{\rho},\mathbf{v},\mathbf{V}) \equiv \{(\rho_t(\omega),v_t,\mathcal{V}_t);\ 0 \le t \le T\}$ is a weak solution of (46), i.e., if

(i) $\boldsymbol{\rho}$ is adapted to $\mathbf{V}$;

(ii) $P\{\omega;\ \int_0^T |\rho_t(\omega)|^2 dt < \infty\} = 1$;

and

(iii) For all $t \in [0,T]$, (46) holds P-a.s. □

Remark 4. Under (C-1) and (C-3), the condition:

$$P\{\omega;\ \int_0^T |\rho_t(\omega)|^2 dt < \infty\} = 1, \tag{48}$$

implies

$$P \times P\{(\omega,\tilde{\omega});\ \int_0^T |h_t(\tilde{\omega}) - \rho_t(\omega)|^2 dt < \infty\} = 1. \tag{49}$$

This is easily seen by

$$\int_0^T |h_t(\tilde{\omega}) - \rho_t(\omega)|^2 dt \le 2\{\int_0^T |h_t(\tilde{\omega})|^2 dt + \int_0^T |\rho_t(\omega)|^2 dt\}, \tag{50}$$

and

$$\int_0^T |h_t(\omega)|^2 dt \le \{\sup_{0 \le t \le T} |H(t)|^2\}\{\int_0^T |f_t(\omega)|^2 dt\} < \infty, \text{ P-a.s.} \tag{51}$$ □

Lemma 1. Assume (C - 1) - (C - 3). Then, $\bar{\mathbf{h}} \equiv \{\bar{h}_t;\ 0 \le t \le T\}$ given by (43) is a strong solution of (46). $\square$

Now, for any $\mathbf{H}$ satisfying (17) and (30), let $\mathbf{S}(\mathbf{H})$ denote a set of nonanticipative functionals defined by

$$\mathbf{S}(\mathbf{H}) \triangleq \{\phi;\ \phi(t,\mathbf{v}),\ 0 \le t \le T \text{ is a strong solution of (46)}\}.$$

By Lemma 1, the proof of which is given below, we see that $\mathbf{S}(\mathbf{H})$ is nonempty.

Lemma 2. Assume (C - 1) - (C - 3). For any solution $\mathbf{y}^*$ of (8) with (13), we have (44) and/or (18) if and only if $\phi \in \mathbf{S}(\mathbf{H})$, i.e., $\phi(t,\mathbf{v}),\ 0 \le t \le T$ is a strong solution of (46). Furthermore, if $\phi \in \mathbf{S}(\mathbf{H})$, then $\hat{\mathbf{h}}^* \equiv \{\hat{h}_t^*;\ 0 \le t \le T\}$ satisfies

$$\hat{h}_t^* = \int_{\tilde{\Omega}} h_t(\tilde{\omega}) \exp\{\int_0^t [h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)]' dy_s^*(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)|^2 ds\} dP(\tilde{\omega}), \quad \text{P-a.s.} \tag{52}$$ $\square$

Now, let us prove Lemma 1 and Lemma 2. For the proof of Lemma 1, we need the following lemma.

Lemma 3. Assume (C - 1) - (C - 3). For any nonanticipative functional ϕ for which (8) has a (weak) solution, denoted by $\mathbf{y}$, let

$$\lambda(t,\mathbf{f},\mathbf{y}) \triangleq \exp\{\int_0^t \beta_s(\mathbf{f},\mathbf{y})' dy_s - \frac{1}{2}\int_0^t |\beta_s(\mathbf{f},\mathbf{y})|^2 ds\}. \tag{53}$$

If

$$P\{\int_0^T |\beta_s(\mathbf{f},\mathbf{y})|^2 ds < \infty\} = 1 \tag{54}$$

and

$$P\{\int_0^T |\beta_s(\mathbf{f},\mathbf{w})|^2 ds < \infty\} = 1, \tag{55}$$

then we have

$$\hat{h}_t \triangleq E\{h_t|\mathcal{Y}_t\} = \frac{\int_{\tilde{\Omega}} h_t(\tilde{\omega})\lambda(t,\mathbf{f}(\tilde{\omega}),\mathbf{y}(\omega))dP(\tilde{\omega})}{\int_{\tilde{\Omega}} \lambda(t,\mathbf{f}(\tilde{\omega}),\mathbf{y}(\omega))dP(\tilde{\omega})}. \tag{56}$$

Proof. See Refs. [12] and [13]. $\square$

Proof of Lemma 1. Let us apply Lemma 3 for the case where $\mathbf{y} = \bar{\mathbf{y}}$ and $\boldsymbol{\beta} = \mathbf{h}$. Then, it is clear by (51) that (54) and (55) hold. Hence, we have

$$\bar{h}_t = \frac{\int_{\tilde{\Omega}} h_t(\tilde{\omega})\lambda(t,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))\,dP(\tilde{\omega})}{\int_{\tilde{\Omega}} \lambda(t,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))\,dP(\tilde{\omega})}. \tag{57}$$

By applying Itô's stochastic differential formula to $\lambda(t,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))$, it can be seen that

$$\begin{aligned}\bar{\lambda}(t,\bar{\mathbf{y}}(\omega)) &\triangleq \int_{\tilde{\Omega}} \lambda(t,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))\,dP(\tilde{\omega}) \\ &= \int_{\tilde{\Omega}} \{1 + \int_0^t \lambda(s,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))h_s'(\tilde{\omega})\,d\bar{y}_s(\omega)\}dP(\tilde{\omega}) \\ &= 1 + \int_0^t \bar{\lambda}(s,\bar{\mathbf{y}}(\omega))\bar{h}_s'(\omega)\,d\bar{y}_s(\omega),\end{aligned} \tag{58}$$

where the third equality follows from (57). Hence, we have

$$\bar{\lambda}(t,\bar{\mathbf{y}}) = \exp\{\int_0^t \bar{h}_s' d\bar{y}_s - \frac{1}{2}\int_0^t |\bar{h}_s|^2 ds\}. \tag{59}$$

Then, it can be seen that

$$\begin{aligned}\frac{\lambda(t,\mathbf{f}(\tilde{\omega}),\bar{\mathbf{y}}(\omega))}{\bar{\lambda}(t,\bar{\mathbf{y}}(\omega))} &= \exp\{\int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' d\bar{y}_s(\omega) - \int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]'\bar{h}_s(\omega)\,ds \\ &\qquad - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \bar{h}_s(\omega)|^2 ds\} \\ &= \exp\{\int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' dv_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \bar{h}_s(\omega)|^2 ds\}.\end{aligned} \tag{60}$$

Hence, it follows from (57) and (60) that

$$\bar{h}_t = \int_{\tilde{\Omega}} h_t(\tilde{\omega})\exp\{\int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' dv_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \bar{h}_s(\omega)|^2 ds\}dP(\tilde{\omega}), \tag{61}$$

which implies that $(\bar{\mathbf{h}},\mathbf{v},\bar{\mathbf{Y}})$, where $\bar{\mathbf{Y}} \triangleq \{\bar{\mathcal{Y}}_t;\ 0 \le t \le T\}$, is a weak solution of (46). Furthermore, we can see that $\bar{\mathbf{h}} \equiv \{\bar{h}_t;\ 0 \le t \le T\}$ is adapted to $\mathbf{V} \equiv \{\mathcal{V}_t;\ 0 \le t \le T\}$ because under (C - 1) - (C - 3), the innovations informational equivalence holds ([9 ; Theorem 1]), i.e.,

$$\bar{\mathcal{Y}}_t = \mathcal{V}_t, \qquad t \in [0,T]. \tag{62}$$

This completes the proof. □

For the proof of Lemma 2, let us prepare the following lemmas.

Lemma 4. Assume (C - 1) - (C - 3). For any $\mathbf{H}$ satisfying (17) and (30), let $\phi \in \mathbf{S}(\mathbf{H})$ and $\mathbf{y}^*$ denote any solution of (8) with (13). Let $\tau_N \equiv \tau_N(\omega)$ be the stopping time defined by

$$\tau_N \triangleq \begin{cases} \inf\{t;\ \int_0^t |h_s(\omega) - \phi(s,\mathbf{y}^*)|^2 ds \geq N \text{ or } \int_0^t |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds \geq N\} \\ T, \text{ if the above set } \{t;\ \cdots\} \text{ is empty.} \end{cases} \tag{63}$$

Then, we have

$$\hat{h}^*_{t\wedge\tau_N} = \phi(t\wedge\tau_N,\mathbf{y}^*) = \int_{\tilde{\Omega}} h_{t\wedge\tau_N(\omega)}(\tilde{\omega})\zeta(t\wedge\tau_N(\omega),\omega,\tilde{\omega})dP(\tilde{\omega}), \quad \text{P-a.s.} \tag{64}$$

and

$$\hat{\zeta}(t\wedge\tau_N,\omega) = 1, \qquad \text{P-a.s.}, \tag{65}$$

where

$$\hat{\zeta}(t,\omega) \triangleq \int_{\tilde{\Omega}} \zeta(t,\omega,\tilde{\omega})dP(\tilde{\omega}) \tag{66}$$

and

$$\zeta(t,\omega,\tilde{\omega}) \triangleq \exp\{\int_0^t [h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*(\omega))]' dy^*_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*(\omega))|^2 ds\}. \tag{67}$$

□

Proof. Let us define $\mathbf{y}^N \equiv \{y^N_t;\ 0 \leq t \leq T\}$ by

$$y^N_t \triangleq \int_0^{t\wedge\tau_N} \{h_s(\omega) - \phi(s,\mathbf{y}^*)\}ds + w_t, \qquad t \in [0,T]. \tag{68}$$

Let $\mathbb{C}^m$ denote the set of continuous functions on [0,T] with values in $\mathbb{R}^m$, and $\mathcal{B}_t$ the σ-algebra generated by all t-cylinder sets[14] in $\mathbb{C}^m$. Also, let $\mu_{\mathbf{y}^N}$ and $\mu_{\mathbf{w}}$ be the distributions of $\mathbf{y}^N$ and $\mathbf{w}$ on $(\mathbb{C}^m,\mathcal{B}_T)$. Then, it can be seen from (68) and (63) that $\mu_{\mathbf{y}^N} \sim \mu_{\mathbf{w}}$, i.e., $\mu_{\mathbf{y}^N}$ and $\mu_{\mathbf{w}}$ are mutually absolutely continuous[10]. Since $\phi(t,\mathbf{v})$ is a strong solution of (46) and $\mu_{\mathbf{y}^N} \sim \mu_{\mathbf{w}}$, we have

$$\phi(t,\boldsymbol{\xi}) = \int_{\tilde{\Omega}} h_t(\tilde{\omega})\exp\{\int_0^t [h_s(\tilde{\omega}) - \phi(s,\boldsymbol{\xi})]' d\xi_s - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \phi(s,\boldsymbol{\xi})|^2 ds\}dP(\tilde{\omega}), \quad \boldsymbol{\xi} \in \mathbb{C}^m, \quad \mu_{\mathbf{w}}\text{- and } \mu_{\mathbf{y}^N}\text{-a.s.} \tag{69}$$

Hence, we have

$$\phi(t,\mathbf{y}^N) = \int_{\tilde{\Omega}} h_t(\tilde{\omega})\exp\{\int_0^t [h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^N(\omega))]' dy^N_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^N(\omega))|^2 ds\}dP(\tilde{\omega}). \tag{70}$$

Since $y_t^N = y_t^*$ for $t \le \tau_N$ and ϕ is nonanticipative, we have

$$\phi(t,\mathbf{y}^N) = \phi(t,\mathbf{y}^*), \quad \text{for } t \le \tau_N. \tag{71}$$

Hence, (70) implies that

$$\phi(t,\mathbf{y}^*) = \int_{\tilde{\Omega}} h_t(\tilde{\omega})\zeta(t,\omega,\tilde{\omega})\,dP(\tilde{\omega}), \quad \text{for } t \le \tau_N, \tag{72}$$

where $\zeta(t,\omega,\tilde{\omega})$ is given by (67). Now, let us apply Lemma 3 for the case $\mathbf{y} = \mathbf{y}^N$ and, hence,

$$\beta_t(\mathbf{f},\mathbf{y}^N) = \begin{cases} h_t(\omega) - \phi(t,\mathbf{y}^N) = h_t(\omega) - \phi(t,\mathbf{y}^*), & \text{for } t \le \tau_N \\ 0, & \text{for } t > \tau_N. \end{cases} \tag{73}$$

Then, by (72) and (66), we have

$$\hat{h}_t^* = \hat{h}_t^N \triangleq E\{h_t(\omega)\,|\mathcal{Y}_t^N\} = \frac{\phi(t,\mathbf{y}^*)}{\hat{\zeta}(t,\omega)}, \quad t \le \tau_N. \tag{74}$$

It can be seen from Itô's stochastic differential formula that for $t \le \tau_N$,

$$\begin{aligned} \hat{\zeta}(t,\omega) &= \int_{\tilde{\Omega}} \zeta(t,\omega,\tilde{\omega})\,dP(\tilde{\omega}) \\ &= \int_{\tilde{\Omega}} \{1 + \int_0^t \zeta(s,\omega,\tilde{\omega})[h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*(\omega))]'dy_s^*(\omega)\}dP(\tilde{\omega}) \\ &= 1 + \int_0^t [1 - \hat{\zeta}(s,\omega)]\phi'(s,\mathbf{y}^*(\omega))\,dy_s^*(\omega), \end{aligned} \tag{75}$$

where the last equality follows from (72). Clearly, (75) implies (65). Also, from (74) and (65), we have (64). This completes the proof. $\square$

Lemma 5. Let $\tilde{\tau}_N \equiv \tilde{\tau}_N(\omega)$ be the stopping time defined by

$$\tilde{\tau}_N \triangleq \begin{cases} \inf\{t;\ \int_0^t |h_s(\omega) - \hat{h}_s^*|^2 ds \ge N \text{ or } \int_0^t |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds \ge N\} \\ T, \quad \text{if the above set } \{t; \cdots\} \text{ is empty.} \end{cases} \tag{76}$$

Then, under the assumptions of Lemma 4, we have

$$\lim_{N\to\infty} \tilde{\tau}_N = T. \tag{77}$$ $\square$

Proof. Note that (77) is implied by

$$P\{\int_0^T |h_s(\omega) - \hat{h}_s^*|^2 ds < \infty\} = 1 \tag{78}$$

and

$$P\{\int_0^T |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds < \infty\} = 1. \tag{79}$$

It can be easily seen from (C-1) and (51) that

$$E\{\int_0^T |h_s(\omega) - \hat{h}_s^*|^2 ds\} \le E\{\int_0^T |h_s(\omega)|^2 ds\} < \infty, \tag{80}$$

which implies (78). Also, since $\mathbf{h}$ and $\mathbf{w}$ are independent, we have

$$P\{\int_0^T |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds < \infty\} = P\times P\{(\omega,\tilde{\omega});\ \int_0^T |h_s(\tilde{\omega}) - \phi(s,\mathbf{v}(\omega))|^2 ds < \infty\}$$
$$= 1, \tag{81}$$

where the last equality follows from $\phi \in \mathbf{S}(\mathbf{H})$ and (49) in Remark 4. This completes the proof. □

Now, let us prove Lemma 2.

Proof of Lemma 2. First, let $\phi \in \mathbf{S}(\mathbf{H})$. Then, by Lemma 4, we have (64). Hence, it can be seen from (63) and (76) that

$$\tau_N = \tilde{\tau}_N, \quad \text{P-a.s.,} \quad \text{for all } N < \infty. \tag{82}$$

By (82) and (77), we have

$$\lim_{N\to\infty} \tau_N = \lim_{N\to\infty} \tilde{\tau}_N = T, \quad \text{P-a.s.} \tag{83}$$

Thus, we have the equalities given by (44) and (52) for $0 \leq t < T$. Also, for $t = T$, (44) and (52) hold because

$$E\{\hat{\zeta}(T,\omega)\} = \int_\Omega \int_{\tilde{\Omega}} \zeta(T,\omega,\tilde{\omega})\, dP(\tilde{\omega})\, dP(\omega) = 1. \tag{84}$$

(Note that by (65) and (83), we have $\hat{\zeta}(t,\omega) = 1$, P-a.s., for $t < T$, and $\hat{\zeta}(t,\omega)$ is left-continuous at $t = T$ by (66).)

Next, let us assume that (44) holds. Then, it can be seen from (41) that $\mathbf{y}^*$ is a Brownian motion process. First, let us show that

$$P\{\int_0^T |h_s(\omega) - \phi(s,\mathbf{y}^*)|^2 ds < \infty\} = 1 \tag{85}$$

and

$$P\{\int_0^T |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds < \infty\} = 1. \tag{86}$$

By (44) and (80), (85) follows. Note that since $\mathbf{y}^*$ is a Brownian motion process, we have

$$E\{|\phi(t,\mathbf{w})|^2\} = E\{|\phi(t,\mathbf{y}^*)|^2\} = E\{|\hat{h}_t^*|^2\} \leq E\{|h_t(\omega)|^2\}. \tag{87}$$

Hence, it follows that

$$E\{\int_0^T |h_s(\omega) - \phi(s,\mathbf{w})|^2 ds\} \leq 2[E\{\int_0^T |h_s(\omega)|^2 ds\} + E\{\int_0^T |\phi(s,\mathbf{w})|^2 ds\}]$$
$$\leq 4\, E\{\int_0^T |h_s(\omega)|^2 ds\} < \infty, \tag{88}$$

which implies (86). Now, let us apply Lemma 3 for the case $\mathbf{y} = \mathbf{y}^*$. Then, we have

$$\phi(t,\mathbf{y}^*) = \hat{h}_t^* = \frac{\int_{\tilde{\Omega}} h_t(\tilde{\omega})\zeta(t,\omega,\tilde{\omega})dP(\tilde{\omega})}{\hat{\zeta}(t,\omega)}, \tag{89}$$

where $\zeta(t,\omega,\tilde{\omega})$ and $\hat{\zeta}(t,\omega)$ are given by (67) and (66). By applying Itô's stochastic differential formula, it can be seen that

$$\begin{aligned}\hat{\zeta}(t,\omega) &= \int_{\tilde{\Omega}} \zeta(t,\omega,\tilde{\omega})dP(\tilde{\omega})\\ &= \int_{\tilde{\Omega}} \{1 + \int_0^t \zeta(s,\omega,\tilde{\omega})[h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*(\omega))]'dy_s^*(\omega)\}dP(\tilde{\omega})\\ &= 1 + \int_0^t [\{\int_{\tilde{\Omega}} h_s(\tilde{\omega})\zeta(s,\omega,\tilde{\omega})dP(\tilde{\omega})\} - \phi(s,\mathbf{y}^*(\omega))\{\int_{\tilde{\Omega}} \zeta(s,\omega,\tilde{\omega})dP(\tilde{\omega})\}]dy_s^*(\omega)\\ &= 1,\end{aligned} \tag{90}$$

where the last equality follows from (89). From (89) and (90), we have

$$\begin{aligned}\phi(t,\mathbf{y}^*) = \int_{\tilde{\Omega}} h_t(\tilde{\omega})\exp\{&\int_0^t [h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*)]'dy_s^*\\ &-\frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \phi(s,\mathbf{y}^*)|^2 ds\}dP(\tilde{\omega}).\end{aligned} \tag{91}$$

Because $\mathbf{y}^*$ is a Brownian motion process, (91) implies that $\phi(t,\mathbf{v})$, $0 \le t \le T$ is a strong solution of (46), i.e., $\phi \in \mathbb{S}(\mathbf{H})$. This completes the proof. □

Now, the proof of Theorem 1 is completed by proving the following lemma.

Lemma 6. Assume (C − 1) − (C − 3). Then, for any $\mathbf{H}$ satisfying (17) and (30), $\bar{\mathbf{h}}$ and $\mathbf{v}$, given by (43) and (42), are respectively the unique strong solutions of (46), and (8) with (13) and (44). That is, for any $\phi \in \mathbb{S}(\mathbf{H})$ and any solution $\mathbf{y}^*$ of (8) with (13), we have

$$\phi(t,\mathbf{y}^*) = \hat{h}_t^* = \bar{h}_t, \qquad \text{P-a.s.}, \qquad t \in [0,T] \tag{92}$$

and

$$\mathbf{y}^* = \mathbf{v}, \qquad \text{P-a.s.} \tag{93}$$ □

In order to prove Lemma 6, let us first show that

Lemma 7. Under the assumptions of Lemma 6, for any $\phi \in \mathbb{S}(\mathbf{H})$ and any solution $\mathbf{y}^*$ of (8) with (13), we have

$$\hat{h}_t^* = \psi(t)\bar{h}_t, \qquad \text{P-a.s.}, \qquad t \in [0,T], \tag{94}$$

where

$$\psi(t) \equiv \psi(t,\omega) \triangleq \exp\{\int_0^t [\bar{h}_s - \hat{h}_s^*]' dy_s^* - \frac{1}{2}\int_0^t |\bar{h}_s - \hat{h}_s^*|^2 ds\}. \qquad (95) \quad \square$$

Proof. By (52) and (61), it suffices to show

$$\exp\{\int_0^t [h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)]' dy_s^*(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)|^2 ds\} = \xi(t,\omega,\tilde{\omega})\psi(t,\omega), \qquad (96)$$

where

$$\xi(t,\omega,\tilde{\omega}) \triangleq \exp\{\int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' dv_s(\omega) - \frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \bar{h}_s(\omega)|^2 ds\}. \qquad (97)$$

Note that

$$\begin{aligned} \int_0^t [h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)]' dy_s^*(\omega) &= \int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' dy_s^*(\omega) + \int_0^t [\bar{h}_s(\omega) - \hat{h}_s^*(\omega)]' dy_s^*(\omega) \\ &= \int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' dv_s(\omega) + \int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' [\bar{h}_s(\omega) - \hat{h}_s^*(\omega)] ds \\ &\quad + \int_0^t [\bar{h}_s(\omega) - \hat{h}_s^*(\omega)]' dy_s^*(\omega), \end{aligned} \qquad (98)$$

where the second equality follows from the fact

$$y_t^* = \bar{y}_t - \int_0^t \hat{h}_s^* ds = v_t + \int_0^t [\bar{h}_s - \hat{h}_s^*] ds. \qquad (99)$$

Also, it can be seen that

$$\begin{aligned} -\frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \hat{h}_s^*(\omega)|^2 ds &= -\frac{1}{2}\int_0^t |[h_s(\tilde{\omega}) - \bar{h}_s(\omega)] + [\bar{h}_s(\omega) - \hat{h}_s^*(\omega)]|^2 ds \\ = -\frac{1}{2}\int_0^t |h_s(\tilde{\omega}) - \bar{h}_s(\omega)|^2 ds &- \int_0^t [h_s(\tilde{\omega}) - \bar{h}_s(\omega)]' [\bar{h}_s(\omega) - \hat{h}_s^*(\omega)] ds \\ &- \frac{1}{2}\int_0^t |\bar{h}_s(\omega) - \hat{h}_s^*(\omega)|^2 ds. \end{aligned} \qquad (100)$$

Then, (98) and (100) implies (96). This completes the proof. $\square$

Proof of Lemma 6. Note that substitution of (94) into (95) yields

$$\psi(t) = \exp\{\int_0^t [1 - \psi(s)] \bar{h}_s' dy_s^* - \frac{1}{2}\int_0^t [1 - \psi(s)]^2 |\bar{h}_s|^2 ds\}. \qquad (101)$$

Let us define $\tilde{\psi}(t)$, $0 \le t \le T$ by

$$\tilde{\psi}(t) \triangleq \psi^{-1}(t) - 1. \qquad (102)$$

Since

$$\psi^{-1}(t) = \exp\{-\int_0^t [1 - \psi(s)] \bar{h}_s' dy_s^* + \frac{1}{2}\int_0^t [1 - \psi(s)]^2 |\bar{h}_s|^2 ds\}, \qquad (103)$$

it follows from Itô's stochastic differential formula that

$$
\begin{aligned}
d\tilde{\psi}(t) &= d\{\psi^{-1}(t)\} \\
&= \psi^{-1}(t)\{-[1-\psi(t)]\bar{h}_t' dy_t^* + [1-\psi(t)]^2 |\bar{h}_t|^2 dt\} \\
&= -\psi^{-1}(t)[1-\psi(t)]\bar{h}_t' dv_t \\
&= -\tilde{\psi}(t)\bar{h}_t' dv_t,
\end{aligned} \tag{104}
$$

where the third equality follows from the fact that

$$
y_t^* = v_t + \int_0^t [\bar{h}_s - \hat{h}_s^*] ds = v_t + \int_0^t [1-\psi(s)]\bar{h}_s ds. \tag{105}
$$

Thus, $\tilde{\psi}(t)$ is the solution of the linear stochastic differential equation:

$$
d\tilde{\psi}(t) = -\tilde{\psi}(t)\bar{h}_t' dv_t, \quad t \in [0,T], \quad \tilde{\psi}(0) = 0. \tag{106}
$$

Hence, we have

$$
\tilde{\psi}(t) = 0, \quad \text{P-a.s.}, \quad t \in [0,T], \tag{107}
$$

or equivalently,

$$
\psi(t) = 1, \quad \text{P-a.s.}, \quad t \in [0,T]. \tag{108}
$$

From (108) and (94), we have (92). Then, it follows that

$$
\begin{aligned}
y_t^* &= \int_0^t [h_s(\omega) - \phi(s,\mathbf{y}^*)] ds + w_t \\
&= \int_0^t [h_s(\omega) - \hat{h}_s^*] ds + w_t \\
&= \int_0^t [h_s(\omega) - \bar{h}_s] ds + w_t \\
&= \bar{y}_t - \int_0^t \bar{h}_s ds = v_t, \qquad \text{P-a.s.}
\end{aligned} \tag{109}
$$

This completes the proof. $\square$

5. Concluding Remarks

In this paper, we were not concerned with the problem of computing the optimal coefficient **H**. Although it seems not so easy as in the case of Gaussian messages to get the optimal coefficient **H**, we can easily find a function **H** which satisfies the constraint on the mean power of the encorded signal. If we can construct a monotone sequence of such coefficients, we have an approximation of the optimal coefficient.

As we can see by the proof of Theorem 1, for the existence, uniqueness and the optimality of functional ϕ with property (18), it is essential that the innovations informational equivalence holds, i.e., $\bar{\mathcal{Y}}_t = \mathcal{V}_t$, $t \in [0,T]$. Because the innovations informational equivalence also holds for the case of the transmission in which the additive noise is a non-Gaussian square integrable martingale[15]-[17], it seems that the result in this paper can be rather easily generalized to this case. The results will be reported in the near future.

Acknowledgements

The author would like to express his sincere thanks to Prof. R. S. Liptser and Prof. P. I. Kitsul of Institute of Control Sciences of the Academy of Sciences, Moscow, USSR for their valuable discussions and comments. Special thanks are extended to Prof. Kurzhanski for his cooperation and encuragement during the study.

REFERENCES

[1] J. P. M. Shalkwijk and T. Kailath, "A cording scheme for additive noise channels with feedback, Part I," IEEE Trans. Inform. Theory, vol. IT-12, pp. 172-182, 1966.

[2] K. Sh. Zigangirov, "Transmission of messages through a binary Gaussian cahnnel with feedback," Problemy Peredachi Informatsii, vol. 3, no. 2, pp. 98-101, 1967.

[3] I. A. Ovseevich, "Optimum transmission of a Gaussian message over a channel with white Gaussian noise and feedback," Problemy Peredachi Informatsii, vol. 6, no. 3, pp. 3-14, 1970.

[4] A. G. Djachkov and M. S. Pinsker, "On optimal linear transmission through a memoryless Gaussian channel with complete feedback," Problemy Peredachi Informatsii, vol. 7, no. 2, pp. 38-46, 1971.

[5] M. B. Nevelson and R. Z. Khasminsky, *Stochastic Approximation and Recursive Estimation*, Nauka, Moscow, 1972.

[6] T. T. Kadota, M. Zakai and J. Ziv, "Mutual information of the white Gaussian channel with and without feedback," IEEE Trans. Inform. Theory, vol. IT-17, no. 4, pp. 368-371, 1977.

[7] S. Ihara, "Cording theory in white Gaussian channel with feedback," J. Multivariate Analysis, vol. 4, pp. 74-87, 1974.

[8] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes II: Applications*, Springer-Verlag, New York, 1978.

[9] D. F. Allinger and S. K. Mitter, "New results on the innovations problem for non-linear filtering," Stochastics, vol. 4, pp. 339-348, 1981.

[10] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes I: General Theory*, Springer-Verlag, New York, 1977.

[11] S. Watanabe, *Stochastic Differential Equations*, Sangyo-tosho, Tokyo, 1975. (In Japanese)

[12] H. Kunita, *Estimation of Stochastic Processes*, Sangyo-tosho, Tokyo, 1976. (In Japanese)

[13] G. Kallianpur and C. Striebel, "Estimation of stochastic systems: arbitrary system process with additive white noise observation errors," Annals of Math. Statistics, vol. 31, no. 3, pp. 785-801, 1968.

[14] Y. Takeuchi and H. Akashi, "Least-squares state estimation of systems with state-dependent observation noise," IFAC J. Automatica, vol. 21, pp. 303-313, 1985.

[15] Y. Takeuchi, "On non-Gaussian innovations processes for observations with non-Gaussian noise," Proc. 25th IEEE Conference on Decision and Control, Athens, Greece, Dec. 10-12, pp. 1035-1036, 1986.

[16] Y. Takeuchi, "Innovations informational equivalence for observations with non-Gaussian noise," Proc. 26th IEEE Conference on Decision and Control, Los Angeles, California, Dec. 9-11, pp. 1163-1168, 1987.

[17] Y. Takeuchi, "Innovationa informational equivalence for a class of observations with independent non-Gaussian noise," (to appear in IEEE Trans. Inform. Theory)